NeuroImage: Clinical 37 (2023) 103290

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Disrupted auditory N1, theta power and coherence suppression to willed speech in people with schizophrenia

Oren Griffiths [a,b,*], Bradley N. Jack [c], Daniel Pearson [d], Ruth Elijah [d], Nathan Mifsud [d], Nathan Han [d], Sol Libesman [d], Ana Rita Barreiros [e,f], Luke Turnbull [a], Ryan Balzan [a,b], Mike Le Pelley [d], Anthony Harris [e,f], Thomas J. Whitford [e]

[a] College of Education, Psychology and Social Work, Flinders University, Adelaide, Australia
[b] Flinders Institute for Mental Health and Wellbeing, Adelaide, Australia
[c] Research School of Psychology, Australian National University, Canberra, Australia
[d] School of Psychology, UNSW Sydney, Sydney, Australia
[e] Specialty of Psychiatry, The University of Sydney, Faculty of Medicine and Health, The University of Sydney, Australia
[f] Brain Dynamics Centre, The Westmead Institute for Medical Research, The University of Sydney, Sydney, Australia

ARTICLE INFO

Keywords:
Schizophrenia
Psychosis
EEG
Electrophysiology
Sensory perception
Sensorimotor control

ABSTRACT

The phenomenon of sensory self-suppression - also known as sensory attenuation - occurs when a person generates a perceptible stimulus (such as a sound) by performing an action (such as speaking). The sensorimotor control system is thought to actively predict and then suppress the vocal sound in the course of speaking, resulting in lowered cortical responsiveness when speaking than when passively listening to an identical sound. It has been hypothesized that auditory hallucinations in schizophrenia result from a reduction in self-suppression due to a disruption of predictive mechanisms required to anticipate and suppress a specific, self-generated sound. It has further been hypothesized that this suppression is evident primarily in theta band activity. Fifty-one people, half of whom had a diagnosis of schizophrenia, were asked to repeatedly utter a single syllable, which was played back to them concurrently over headphones while EEG was continuously recorded. In other conditions, recordings of the same spoken syllables were played back to participants while they passively listened, or were played back with their onsets preceded by a visual cue. All participants experienced these conditions with their voice artificially shifted in pitch and also with their unaltered voice. Suppression was measured using event-related potentials (N1 component), theta phase coherence and power. We found that suppression was generally reduced on all metrics in the patient sample, and when voice alteration was applied. We additionally observed reduced theta coherence and power in the patient sample across all conditions. Visual cueing affected theta coherence only. In aggregate, the results suggest that sensory self-suppression of theta power and coherence is disrupted in schizophrenia.

1. Introduction

When a person generates a sound, the neural processing of that sound is reduced relative to the processing of an otherwise identical sound that they did not produce (Hughes et al., 2013). This "sensory self-suppression" effect has been demonstrated in several species using a broad range of experimental methodologies (Eliades and Wang, 2003). In humans it has been documented across vision and audition (see Bansal et al., 2018), but is most commonly studied in auditory-evoked response potentials (ERPs) measured by electroencephalography (EEG). In a typical study, a person first generates a sound by speaking, and their own voice is synchronously played back to them over headphones (Ford et al., 2010). This is the Active condition. Neural responses to the sound in this Active condition are compared against a second Passive condition in which the participant listens to a playback of that same voice recording without undertaking any action. The neural index of primary auditory processing, commonly the N1 in EEG (Näätänen and Picton, 1987) or the M1 in magnetoencephalography (Roberts et al., 2000) is reduced in the Active condition relative to the Passive condition. This reduction in cortical responsiveness – referred to as sensory

* Corresponding author at: College of Education, Psychology & Social Work, Sturt Road, Bedford Park, South Australia 5042, Australia.
E-mail address: oren.griffiths@flinders.edu.au (O. Griffiths).

https://doi.org/10.1016/j.nicl.2022.103290
Received 8 February 2022; Received in revised form 17 November 2022; Accepted 12 December 2022
Available online 13 December 2022

self-suppression – is notable because the auditory stimulus is identical in the Active and Passive conditions. Of particular importance is that this sensory self-suppression (difference between Active and Passive conditions) is reliably reduced in people with a diagnosis of schizophrenia (for reviews see Whitford, 2019; Ford and Mathalon, 2012), bipolar disorder (Ford et al., 2013), people who score highly on the schizotypy personality dimension (Oestreich et al., 2015), or people who are at heightened levels of risk for psychosis (Mathalon et al., 2019).

Research examining the relationship between sensory self suppression and specific symptoms has produced mixed findings. In a recent review, Whitford (2019) examined 19 experiments assessing sensory self-suppression in people with psychosis versus controls. Of the 14 studies identified by Whitford (2019): 2 did not look at relationships with specific clinical symptoms, 8 found no evidence of a relationship with clinical symptoms, 1 found evidence on a secondary measure, and 3 found positive evidence of a relationship with different clinical symptoms. Work in this area has historically been challenged by low sample sizes and variation between symptom measures, and variance within the measures themselves. More recently, however, Mathalon et al. (2019) reported findings from a much larger sample of people with schizophrenia (84 people), people at high risk of developing schizophrenia (71) and healthy controls (103). They observed a clear reduction in suppression (relative to controls) in both people with a diagnosis of schizophrenia, and those at high risk of the condition. With respect to particular symptoms, they observed a relationship between N1 suppression magnitude and unusual thought content in the high risk group, but this was not replicated in the patient sample (no symptom clusters were associated with specific symptoms in the patient sample).

An alternative approach to examine sensory self-suppression is to use spectral decomposition. This allows investigators to refine investigations of the perturbation of ongoing neural activity brought about by sensory prediction, and to examine any disruption in this process associated with clinical conditions. Spectral analyses of this kind have the additional benefit of being relatively less sensitive than event-related potential metrics to some forms of artifacts (broadband noise) that differ between individuals, and could conceivably systematically differ between patients and controls (movement artifacts, differences in general alertness indexed by alpha activity, and so on; Cohen, 2014). Under ordinary passive listening conditions in adults, electrophysiological responses to auditory stimuli are characterized by transient, phase-locked increases in theta (4–7 Hz) power, typically observed broadly between 0 and 300 ms post stimulus onset (Yordanova and Kolev, 1998; Kolev et al., 2001). Theta power has been shown to be altered in people with schizophrenia during passive listening tasks (Basar-Eroglu et al., 2008) and while at rest (Andreou et al., 2015; König et al., 2001).

Roach et al. (2021) explored whether sensory self-suppression was evident in a reduced phase-locked theta response, and whether this suppression was itself reduced in those with a diagnosis of schizophrenia (see also (Ford et al., 2002), for an earlier investigation). It was important to investigate both theta power and phase (coherence), as there is increasing evidence that sensory event-related potentials (like the auditory N1) may not just be the product of transient increases or decreases in electrical activity, but may instead be primarily the product of phase resetting of ongoing endogenous oscillations (Sauseng et al., 2007; Fuentemilla et al., 2006).

Roach et al. (2021) found that suppression was evident not just in traditional event-related potential measures (the N1 component), but also in theta intertrial coherence and evoked theta power. Indeed, they found evidence that coherence (specifically, intertrial coherence, measured as amplitude-normalized phase angle) was the most sensitive measure of suppression, with an effect size approximately double that of suppression measured by N1 magnitude or event-related power. Moreover, they found a relationship between suppression of theta coherence and delusion severity within the patient sample, even after controlling for magnitude of suppression of the N1 response. Interestingly, these predictive properties appeared somewhat specific to the coherence metric, and reductions in event-related theta power were significantly weaker than suppression of coherence. Cumulatively, these findings were interpreted as evidence that suppression of theta coherence may be a specific biomarker for schizophrenia, as distinct from traditional N1 suppression which may index a more general psychosis vulnerability. The N1 component is a complex amalgam of several sources (Woods, 1995) and contains both spontaneous (or nonphase-locked) power in the theta range, and power from other frequency bands (Grau et al., 2007). In contrast, the theta coherence measure is specific to phase-resetting activity, which is argued to be initiated by the detection of an external sound stimulus and suppressed by the presence of a concurrent motor command to initiate speech (Wang et al., 2014). Further, because the intertrial coherence metric is normalized by amplitude, it is less directly impacted by any gross group differences in the theta amplitude. This property may partly explain why Roach et al found that suppression of intertrial coherence, but not of power or N1 magnitude, was able to discriminate people with schizophrenia from both a healthy control sample and group at clinical high risk (when controlling for covariation). This prompted Roach et al. (2021) to posit that theta coherence may be an improved biomarker of schizophrenia specifically.

The present research sought to further examine this claim by testing the effect of stimulus predictability on the same three metrics (N1 amplitude, theta coherence and theta power) in both a patient sample and healthy control sample. N1 suppression is affected by predictability. Prediction error processing is known to be disrupted in schizophrenia (Millard et al., 2022; Fletcher and Frith, 2009) , N1 amplitude is reduced in schizophrenia (Mathalon et al., 2019), and N1 amplitude is affected by the predictability of an incoming sound stimulus (Bendixen et al., 2012). More specifically, the N1 response to speech appears to be affected by both the temporal predictability of the incoming sound stimulus (Van Wassenhove et al., 2005) and the predictability of its content (Behroozmand and Larson, 2011). The broad concordance of these effects has led some theorists (Whitford, 2019; Ford and Mathalon, 2012) to argue that the N1 may be a particularly important marker for the cognitive disruptions that give rise to hallucinations (and/or delusions). It is therefore important to examine whether theta coherence and power retain these same properties. To this end, the present experiment modulated the predictability of the sound playback in two distinct ways. Temporal predictability was manipulated by providing a visual cue as to the onset of the sound. Second, on half of the trials, participants' voice playback was shifted up by 5 semitones. We predicted that an increase in temporal predictability would reduce N1 magnitude, while a decrease in the predictability of the content (via pitch-shifting) should increase N1 magnitude. It was expected that both of these modulations would be more pronounced in the healthy sample than the patient sample, given the putative reduction in predictive processing in people with schizophrenia (see also Heinks-Maldonado et al., 2007). Finally, because theta power had less sensitivity to diagnostic status in Roach et al.'s (2021) study, we sought to quantify theta power differently, and potentially more sensitively as we could capitalize on Roach et al. (2021) findings to tailor our own approach. Specifically, instead of deriving theta power from general purpose Morlet wavelet transformations, we used a filter-Hilbert (Cohen, 2014) approach. Under this approach, a narrowband filter focused on the theta band is first applied, and then the Hilbert transform is used to determine the evolution of theta power across time. This custom approach allows more power to quantify a particular signal of interest (at the cost of reducing meaningful comparison across different frequencies).

2. Method

2.1. Participants

The experiment was approved by the Western Sydney Local Health District Human Research Ethics Committee, and followed the Code of Ethics of the World Medical Association (Declaration of Helsinki). We

recruited as many people with psychosis as possible across 1.5 years. Sixty-one participants were recruited in total: 33 people with a diagnosis of schizophrenia and 28 controls. Each clinical participant was referred by their treating psychiatrist at Cumberland Hospital (Sydney, Australia). Every participant signed a consent form prior to enrolment in the study. In the clinical group, diagnosis was confirmed using the Structured Clinical Interview for DSM IV (SCID; First et al., 1997); measures of symptom severity were collected (Positive and Negative Symptom Scale, Kirkpatrick et al., 2011); Psychosis Rating Scale, Haddock et al., 1999); screens for post-traumatic stress disorder (PSTD), substance and alcohol abuse were performed using the Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998); premorbid intelligence was estimated using the National Adult Reading Test (NART; McGurn et al., 2004), global functioning was measured using the Personal and Social Performance scale (Nasrallah et al., 2008), and self-reported mood was collected using the Depression, Anxiety and Stress Scale (DASS-21; Lovibond and Lovibond, 1995). Control participants were screened for any current mood disorder, lifetime PTSD, alcohol or substance abuse or psychotic disorders. Of the initial 61 participants, only 26 patients and 26 controls completed the experiment (see Results for details) and were thus included in the final sample. The clinical group consisted of 22 males and 4 females (mean age = 33.31 years, SD = 13.10). They had experienced an average of 11.16 years of illness (SD = 10.37 years) and were medicated on a mean dose of 767.50 mg/day of chlorpromazine equivalent antipsychotic medication (SD = 1101.85 mg/day). The control group consisted of 17 males and 9 females (mean age = 34.19 years, SD = 14.48). Further descriptive data for the clinical group is summarized in Table 1.

2.2. Apparatus

Participants spoke into a Shure M52 microphone, which was routed through a low-latency pitch-shifting device (Digitech Whammy 4 pitch shifter), and then to the stimulus delivery computer (using a low-latency, ASIO compliant audio card), and on to a set of over-ear headphones (Sennheiser HD30) worn by the participant. Volume was controlled such that volume during speaking did not exceed ~ 80 dB SPL. The pitch shifter was controlled by the stimulus computer (via a Musical Instrument Digital Interface port) and could be enabled or disabled on any given trial. A second speech signal was split off and sent directly to the EEG amplifier so that auditory signals could be digitized alongside EEG signals with high temporal precision. Audio and visual stimulus delivery, audio recording and playback was controlled by a custom script written for Matlab (MATLAB, V., 2010) calling the Psychophysics Toolbox (Kleiner et al., 2007) package. EEG was recorded with a BioSemi ActiveTwo 64-channel system, using Ag/AgCl active electrodes placed according to the extended 10–20 system.

Table 1
Descriptive data for the clinical group derived from symptom measures and screening protocols. See Procedure for more detail.

Symptom Measure	Mean (SD)
DASS – Depression	12.88 (12.25)
DASS – Anxiety	14.25 (11.73)
DASS – Stress	15.38 (13.22)
PSYRATS- Hallucinations	19.71 (16.23)
PSYRATS – Delusion	10.45 (8.18)
PANSS – Positive	18.16 (6.02)
PANSS – Negative	19.53 (7.04)
PANSS – General	35.88 (8.99)
NART – IQ est.	109.07 (9.10)
PSP	51.42 (14.42)
Screening measure	**Met criteria**
MINI- Substance Abuse	3/26
MINI – Substance dependent	6/26
MINI – Alcohol Abuse	1/26
MINI – Alcohol Dependent	6/26
MINI – PTSD	1/26

2.3. Design

The experiment contained one between-subjects pseudo-experimental factor (Group) and two repeated-measures factors (Action and Voice). With regard to the Action factor, all participants generated a series of vocal sounds in the Talk condition, and then listened to these same sounds passively in the Listen and Cued conditions, respectively. With regard to the Voice factor, participants performed all Action conditions with their regular, unshifted vocal utterances (Normal condition), and with their utterances shifted 5 semitones higher (Shift condition).

2.4. Procedure

Participants sat in a quiet room, 60 cm in front of a computer monitor (24″, 1920 × 1080 resolution) and EEG electrodes were applied. EEG was recorded continuously throughout the task. A vertical electro-oculogram was calculated by recording from an electrode placed below the left eye, and subtracting its activity from that of electrode FP1; a horizontal EOG was recorded by placing an electrode on the outer canthus of each eye. We also placed an electrode on the tip of the nose, on the left and right mastoid, and on the masseter muscle to detect jaw movements. During data acquisition, the reference was composed of CMS and DRL sites, and the sampling rate was 2048 Hz.

Following electrode placement, participants were instructed on how to perform during the Talk condition – to make an "ah" sound ("a" in the international phonetic alphabet; Duckworth et al., 1990) every 2–4 s, and to otherwise remain as still as possible. A video of the experimenters performing the task was shown to facilitate instruction. Ten practice utterances were then performed. Auditory stimulus onsets during the task were calculated using a voice trigger, whereby the stimulus computer sent a trigger to the EEG data stream whenever an utterance louder than a fixed input level was recorded (implemented as custom MATLAB script). The practice utterances were used to calibrate the appropriate input level such that vocalizations elicited triggers but ambient noises (e.g. movements in the chair) did not. The Normal Talk condition commenced after the practice/calibration task. Participants watched a white fixation cross on a black screen, which did not change across the condition. They then spoke until they had recorded 60 distinct utterances (all vocalizations were recorded for subsequent playback, but the voice trigger would not record a new utterance until 500 ms after the preceding utterance).

Next, participants completed the Normal Listen condition. They were asked to listen passively while a recording of their voice from the Talk condition was played back to them. The screen was identical to the Normal Talk condition. Once this was complete, participants were asked to listen to their voice recording one further time in the Cued condition. As they listened they watched a green trace of their voice (the time series of their audio recording) scroll from right to left across the screen. A stationary red cursor placed in the centre of the screen indicated the current moment of playback. In this manner people could use the visual information to predict the onset of each utterance. Participants then completed the same 3 conditions (Talk, Listen, Cued) in the same sequence, but their voice was shifted 5 semitones higher. In all other respects, the Shift conditions were identical to the Normal voice conditions. Following completion of all experimental conditions and a 1 h break, a clinical interview was administered and self-report questions were completed.

2.5. EEG analysis

Data were processed using MATLAB in conjunction with two toolboxes: EEGlab (Delorme and Makeig, 2004) and PREP pipeline (Bigdely-

Shamlo et al., 2015). EEG data were first down-sampled to 256 Hz, and then subjected to the default parameters of the PREP pipeline. Drift artefacts were minimized by DC offset, before line noise was removed using the CleanLine multi-taper algorithm (Mullen, 2012). Data were then band-pass filtered from 0.1 to 30 Hz using a phase-shift free Butterworth filter (fourth order), before being referenced to the linked mastoids. Post hoc analyses revealed that channel FCz was affected by a small but reliable artefact. We assessed this as being likely due to pressure of the headphone band on electrodes in the frontocentral band, particularly at the midline. Specifically, it yielded measurements approximately 1 μV higher than surrounding channels, despite being correlated $>$ 0.9 with neighbouring channels. FCz was interpolated using EEGlab's default spherical interpolation algorithm.

ICA was used to identify eye-blink and horizontal eye movement artefacts (implemented as Matlab's "runica" function), with a mean of 1.95 components removed per participant (range = 1–3). Data were then epoched to the onset of each vocal utterance. Epochs started 1000 ms prior to stimulus onset, and terminated 1000 ms after stimulus onset. Each epoch was baseline-corrected to the average voltage 100 ms prior to stimulus onset. Artefact rejection applied a 100uV simple voltage threshold applied to all scalp channels, and resulted in the removal of 10.64 % of all epochs. This left an average of 46.88 epochs per participant, per condition. To maintain consistency across analyses, listwise deletion was used for partial completers (26 people in each sample completed the task). One further clinical participant's data were removed due to theta power measures (both during the task and during baseline) that were 3 orders of magnitude greater than the remaining sample. This left 26 participants in the control group, and 25 in the patient group.

The amplitude of the N1-component of the auditory-evoked potential is typically measured at the midline centrally or frontocentrally, approximately 100 ms after stimulus onset (Näätänen and Picton, 1987). A collapsed localizer (collapsing across groups and all conditions) was used to find the time of the most negative voltage point within a broad montage of nine frontocentral sites centered on FCz (FCz, FC1, FC2, Cz, C1, C2, Fz, F1, F2), as recommended by (Luck and Gaspelin, 2017). The N1 peak was observed approximately 110 ms after stimulus onset (113.28 ms). Mean amplitude during a 50 ms window (85–135 ms) was used to quantify the N1. Sensory self-suppression appears to be driven by two N1 sub-components (N1b and N1c) and is typically maximal at midline frontocentral or central sites (Timm et al., 2013).

Phase coherence and theta power were both quantified from the same epoched data used to calculate ERPs. Phase (intertrial) coherence was estimated by submitting the single-trial, epoched data to Morlet wavelet decomposition (using the default "newtimef" values in EEGlab, except where specified). The window size was increased to 143 samples (0.56 s) and the wavelets commenced at 1 cycle at 2 Hz increasing linearly (0.5cycles/Hz) through to 50 Hz. Baseline correction was applied per epoch (1 s prior to sound onset). Phase coherence was quantified as intertrial coherence (or "phase locking factor") measured as amplitude normalized mean phase angle. ITC was quantified as the mean across bins within 4–7 Hz, and within a 100 ms window centred on N1 peak amplitude (60–160 ms post sound onset). Inferential analyses were performed on the average of all 64 scalp channels.

Theta power was quantified using the filter-Hilbert method (Cohen, 2014). The same 2 s epochs were narrow-band filtered to 4–7 Hz (using a Guassian filter centered on 5.5hz with 3 Hz full-width at half maximum; see Cohen and Gulbinaite, 2017), then subjected to the Hilbert transformation. Signal to Noise Ratio (SNR) in decibels was calculated by calculating power at each moment during the epoch (the square of the magnitude of the complex Hilbert output), dividing this by the mean power during the baseline period (500 to 0 ms pre stimulus onset), and taking 10 times the logarithm (base 10) of the resulting value. Theta power was quantified as mean SNR from 0 to 250 ms after stimulus onset and inferential analyses were performed on the mean of all channels.

3. Results

Event-related potentials for each group and condition are shown in Fig. 1. N1 deflections were clearly evident in both groups and in all conditions, albeit substantially reduced in the Talk conditions. This suppression was evident in each group and condition, but generally appeared reduced in the patient sample, and to a lesser extent in the pitch-shifted voice condition. Spectral measures of theta power showed a similar general pattern (as seen in Fig. 2), although suppression looked substantially reduced in the patient sample on this measure. Coherence measures shown in Fig. 3 showed clear evidence of a theta-specific period of phase coherence approximately synchronous with N1 peak magnitude, and with a broad frontocentral scalp distribution (see Fig. 3). Coherence appeared to be absent during the Talk condition, and markedly reduced in the patient sample relative to controls.

All dependent variables (N1 magnitude, phase coherence, and theta power) were analyzed using 2 × 2 × 3 ANOVA, with one between-subjects measure (group: schizophrenia or control), and 2 repeated measures factors, Action (Talk, Listen, Cued) and Voice (shifted, normal). Within the Action factor, two planned, orthogonal contrasts were tested (as custom contrasts in JASP; JASP Team, 2021). The first compared the active (Talk) condition against the average of the two passive conditions (Cued, Listen). The second compared the two passive conditions to test the effect of visual cueing (Cued versus Listen). The obtained values entered into each analysis are summarized in Fig. 4.

N1 magnitude. As shown in Fig. 1, N1 suppression appeared to be evident to some extent in all groups and conditions. The ANOVA revealed no main effect of group, $F(1, 49) = 2.44$, $p = .13$, $\eta_P 2 = 0.05$, nor was there an effect of Voice, $F(1, 49) = 0.81$, $p = .37$, $\eta_P 2 = 0.02$. Of the two main effect contrasts testing the Action factor, N1 magnitude was significantly larger (more negative) in the two passive conditions than the active (Talk) condition, $t(98) = 6.50$, $p < .001$, $d = 0.66$, but there were no differences between the Listen and Cued conditions, $t(98) = 0.41$, $p = .68$, $d = 0.04$. The comparison between active and passive conditions significantly interacted with Voice, $t(98) = 2.42$, $p = .02$, $d = 0.24$, the interaction with group approached significance, $t(98) = 1.87$, $p = .06$, $d = 0.19$, but the three-way interaction was not significant, $t(98) = 1.62$, $p = .11$, $d = 0.16$. Follow-up tests indicated that N1 magnitude was significantly larger (more negative) in the passive than active conditions in all groups and Voice conditions when considered individually: patients in normal voice condition, $t(24) = 2.39$, $p = .03$, $d = 0.48$, patients in pitch-shifted condition, $t(24) = 2.79$, $p = .01$, $d = 0.56$, controls in normal voice condition, $t(25) = 3.88$, $p < .001$, $d = 0.76$, and control in pitch-shifted condition, $t(25) = 3.51$, $p = .002$, $d = 0.69$. In summary, N1 suppression was evident in both groups and in both Voice conditions individually, but there was evidence that the degree of suppression was larger in the unaltered voice condition than the pitch-shift condition, and a marginally significant effect indicating greater suppression in the control group than in patients.

3.1. Phase coherence

Turning to phase coherence, there were significant main effects of group, $F(1, 49) = 22.32$, $p = .001$, $\eta_P 2 = 0.31$, and Voice, $F(1, 49) = 5.13$, $p = .03$, $\eta_P 2 = 0.10$. Phase coherence was greater in the control sample and in the unaltered voice condition, respectively. The main effect contrast testing active (Talk) versus passive (Listen, Cued) conditions, was significant, $t(98) = 5.31$, $p < .001$, $d = 0.55$. This contrast did not interact with Voice, $t(98) = 0.58$, $p = .56$, $d = 0.06$, there was a marginally significant interaction with group, $t(98) = 1.95$, $p = .05$, $d = 0.20$, and the three-way interaction was not significant, $t(98) = 0.14$, $p = .89$, $d = 0.01$. Follow-up t-tests revealed suppression effects (Talk versus Listen and Cued) in the patient group: pitch-shifted condition, $t(24) = 2.09$, $p = .048$, $d = 0.42$, and a marginally significant effect in the unaltered condition, $t(24) = 1.96$, $p = .06$, $d = 0.39$. In the control group, suppression was evident in the pitch-shifted condition, $t(25) = 3.02$, p

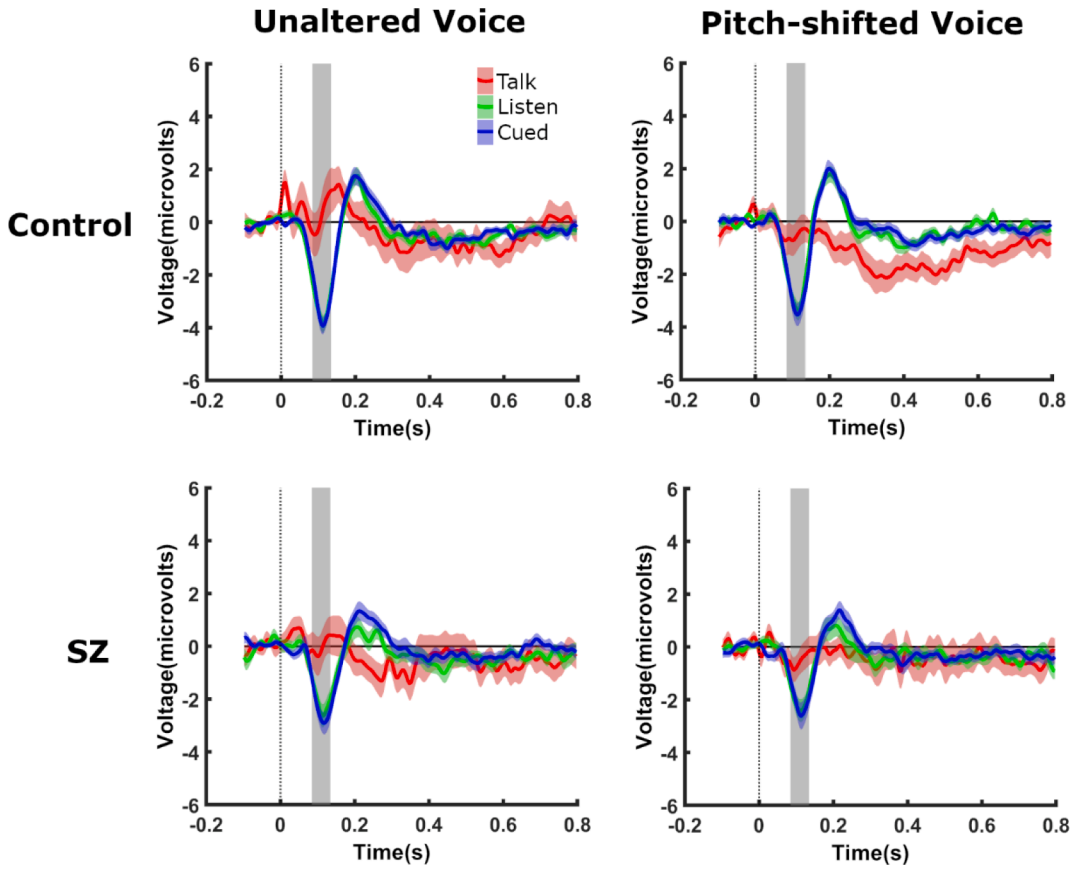

Fig. 1. Event-related potentials across conditions and groups, measured in microvolts. The upper row shows the performance of the control sample, and the lower row shows the performance of the patient sample (labelled SZ). The left-hand column shows performance on the unaltered voice trials, the right-hand column shows performance on the pitch-shifted trials. The grey shaded area shows the time window used for N1 quantification. Each coloured line depicts mean performance, and the shaded area represents the standard error of the mean.

=.006, $d = 0.59$, and the unaltered condition, $t(25) = 3.28$, p =.003, d = 0.64.

Turning to the effect of cueing, the main effect contrast testing the effect of cueing was significant, $t(98) = 2.09$, p =.04, $d = 0.20$, with cueing eliciting greater intertrial coherence. This main effect did not significantly interact with Voice condition, $t(98) = 0.22$, p =.83, d = 0.01, nor group, $t(98) = 0.32$, p =.75, $d = 0.03$, nor both factors simultaneously, $t(98) = 0.58$, p =.56, $d = 0.06$.

In summary, phase coherence was reduced in the patient sample relative to controls, and in the voice-altered condition relative to the unshifted condition. With respect to the effect of cueing, we observed a main effect that did not interact with pitch-shifting or group membership. With respect to suppression, suppression was generally evident across groups and all conditions (albeit one simple effect was marginal), but there was a trend towards phase coherence being larger in controls than in the patient group.

3.2. Theta power

There was a main effect of group, $F(1, 49) = 15.79$, p <.001, $\eta_P 2$ = 0.24, but not of Voice condition, $F(1, 49) = 0.03$, p =.85, $\eta_P 2 < 0.001$. Control participants showed a larger increase in theta power than patients in response to sound onset. The main effect of Action testing suppression (Talk versus Listen and Cued) was significant, $t(98) = 7.74$, p <.001, $d = 0.79$. This contrast significantly interact with group, $t(98)$ = 2.52, p =.01, $d = 0.26$, there was a marginally significant interaction with Voice, $t(98) = 1.69$, p =.09, $d = 0.17$, and no significant three-way interaction, $t(98) = 0.16$, p =.87, $d = 0.01$. Follow up t-tests revealed that the talk condition yielded significantly less evoked theta power than the listen/cued conditions in the unaltered voice condition in the control sample, $t(24) = 5.04$, p <.001, $d = 0.99$, the pitch-shifted voice condition in the control sample, $t(24) = 3.66$, p =.001, $d = 0.72$, the unaltered voice condition in the patient sample, $t(25) = 2.26$, p =.03, $d = 0.45$, but not in the pitch-shifted voice condition in the patient sample, $t(25)$ = 1.06, p =.30, $d = 0.21$.

With respect to the effect of cueing, there was no main effect of cueing (Listen versus Cued), $t(98) = 0.34$, p =.74, $d = 0.03$. This main effect contrast did not significantly interact with Voice, $t(98) = 0.72$, p =.48, $d = 0.07$, nor with group, $t(98) = 0.67$, p =.50, $d = 0.07$, nor with their interaction, $t(98) = 0.04$, p =.97, $d = 0.003$.

In summary, theta power evoked by sound onset was lower across all conditions in the patient sample than the control sample. There was evidence of N1 suppression of theta power, but this was larger in the control sample than the patient sample, and reduced in magnitude with voice alteration. No effect of cueing on theta power was observed.

4. Discussion

Consistent with earlier studies of sensory suppression in patients with schizophrenia (Ford et al., 2001; Ford et al., 2001), N1 magnitude was reliably lower when people were speaking than when they were passively listening. Further, sound-evoked responding produced clear, widespread and spectrally specific phase-coherence in the theta band. Theta coherence and power were suppressed in the Talk condition relative to the two passive conditions (Listen and Cued), replicating Roach et al. (2021). The magnitude of suppression was comparable across all metrics, and if anything was larger in our power metric: N1 magnitude ($d = 0.66$), phase coherence ($d = 0.55$), and theta power (d =

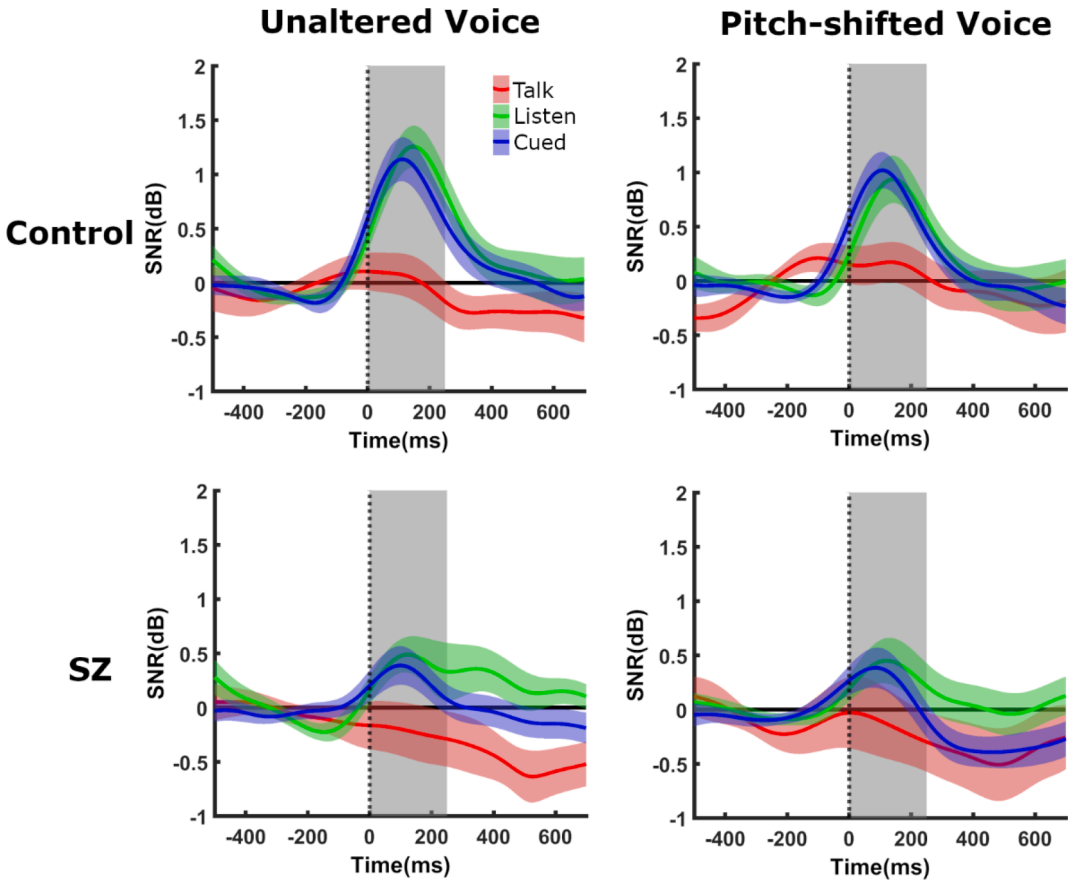

Fig. 2. Mean baseline-corrected theta power across conditions and groups, measured as signal-to-noise-ratio (SNR) in decibels averaged across all 64 scalp channels. The upper row shows the performance of the control sample, and the lower row shows the performance of the patient sample (labelled SZ). The left-hand column shows performance on the unaltered voice trials, the right-hand column shows performance on the pitch-shifted trials. The grey shaded area shows the time window used for theta power quantification. Each coloured line depicts mean performance, and the shaded area represents the standard error of the mean.

0.79). While this is distinct from the pattern seen in Roach et al. (2021), it is premature to interpret the two findings as in conflict. There are potentially important differences between the two. Roach et al. (2021) tested a larger sample, primarily comprising patients during early stages of the illness (our sample included a wide spectrum of illness presentations, with a mean of 10 years post onset). There were also potentially important distinctions in postprocessing: Roach et al analyzed principal components derived from their spectral analyses, while we report the baseline-corrected time–frequency data directly.

Despite these sample and procedural differences, it was encouraging to observe that our study produced broadly similar findings to Roach et al. (2021) with respect to suppression. Sound onset produced a spectrally specific response in the theta band, which was characterized by phase alignment peaking at a comparable time to the N1 event-related potential response. All three metrics (N1 amplitude, theta coherence and theta power) showed action-related suppression in both experiments. Both studies showed evidence that healthy controls differed from people with schizophrenia in suppression metrics, but the broad pattern differed. Roach et al. (2021) found significant differences between patients and controls when aggregating across all measures, and noted that, if anything, coherence was the clearest discriminator between patients and controls. By contrast, the interactions with group membership in the present study were marginal for N1 magnitude ($p = .06$) and coherence ($p = .05$), and instead only the interaction involving power was statistically significant ($p = .01$). Note further that the direct comparison across methods in the current study versus Roach et al. is additionally challenged by model terms. For example, our statistical model included Cued and Uncued listening conditions; Roach et al. included a single Listen condition (which used a different kind of visual display to that deployed here). Nevertheless, the present data provide convergent evidence that closer examination of theta-band activity during speech processing may selectively differentiate people with schizophrenia from those without that diagnosis.

Indeed, our detection of suppression differences between patients and controls in theta activity was challenged by the presence of large main effect differences in theta coherence and power between the two groups. People with schizophrenia showed less event-related changes in theta power and phase alignment than controls when averaged across active and passive conditions. However, from Fig. 4 it appears this difference is primarily evident in differential responses to speech in the listening condition, rather than during the active performance of speech. This general pattern was also evident in Roach et al. (2021) findings. It appears that people with schizophrenia have a reduced theta response when listening to speech. This is broadly consistent with observations of schizophrenia impacting a broad range of neural oscillations that have been linked with cognition (see Uhlhaas and Singer, 2010 for a review).

Interestingly, the theta-specific response profile does not appear to generalize to all sounds. In the mismatch negativity (MMN) paradigm, which typically uses simple sinusoidal tones, Javitt et al. (2000) showed that the spectral response profile to a standard tone involved broad power over the theta and alpha range. By contrast, the response to deviant stimuli (infrequent stimuli that typically differ in one characteristic from the frequent, standard tone) was followed by a profile that more clearly emphasized the event-related increase in theta power. It was this theta-dominated response to deviant stimuli that primarily differentiated between people with schizophrenia and those without. In a more recent examination of this effect, Lee et al. (2017) demonstrated reduced power and coherence in the theta band during deviant stimuli in

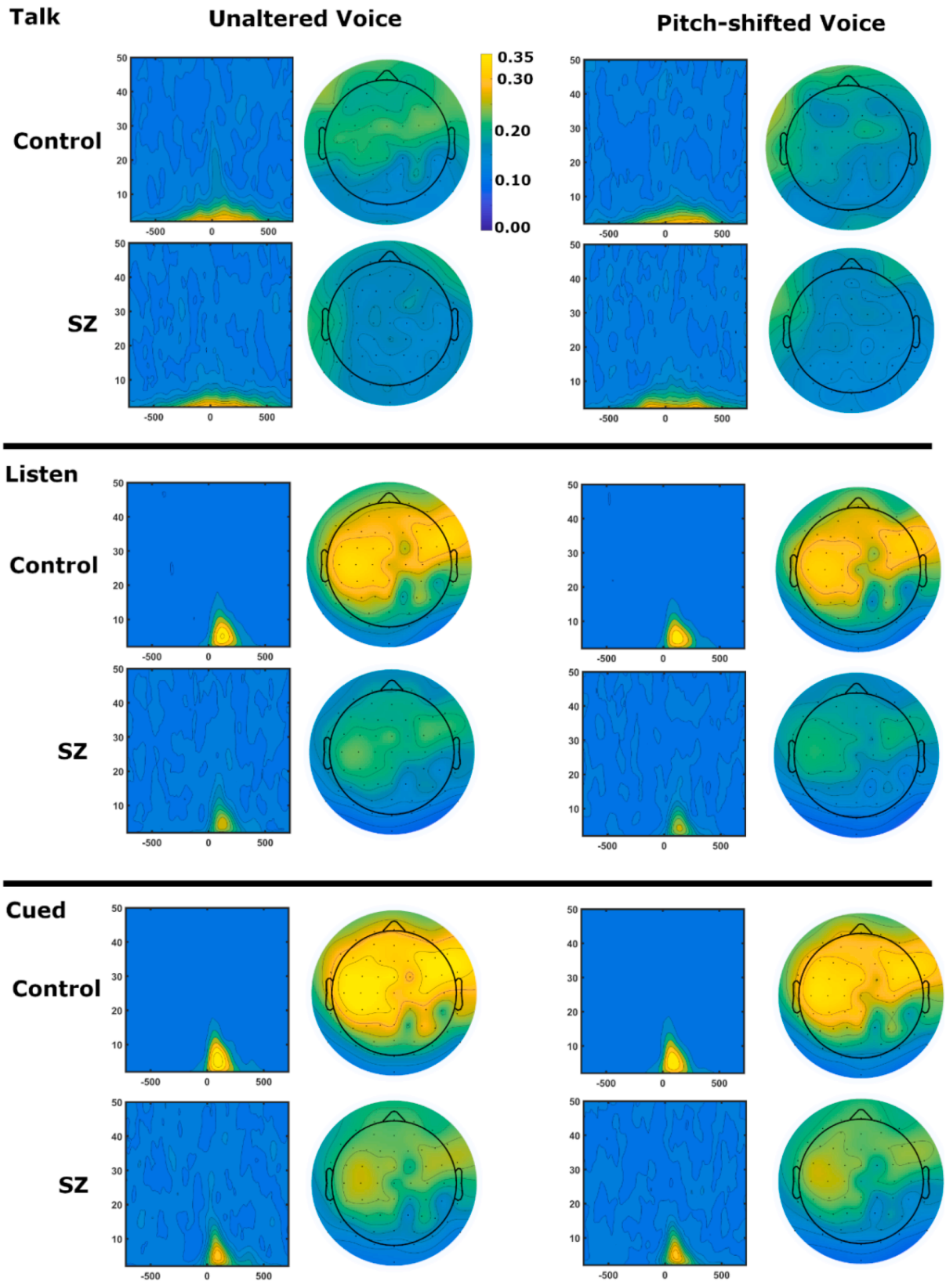

Fig. 3. Intertrial coherence across conditions and groups. The upper, middle and lower sections depict intertrial coherence in the Talk, Listen and Cued conditions, respectively. Within each section the left-hand images depict coherence during the unaltered voice trials, and the right hand images show coherence during pitch-shifted trials. Within each section, the upper row depicts the performance of the control group, and the lower row depicts performance of the patient group (labelled SZ). Time frequency plots show coherence distributed across time (-700 to 700 ms) and frequency (2–50 Hz), and averaged across all 64 scalp channels. The topoplots show mean coherence across at each channel, averaged across theta (4–7 Hz) and between 60 and 160 ms after stimulus onset. All panels used the same colour axis, which refers to the degree to which measured amplitudes are phase aligned across trials (0–1).

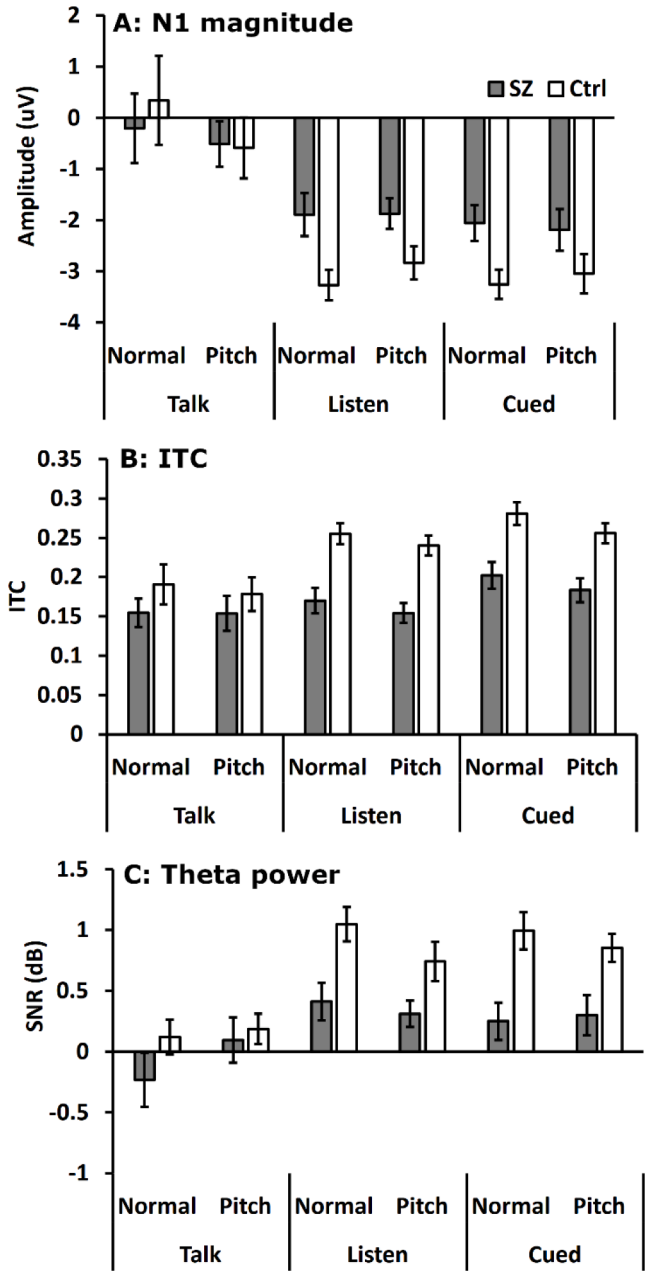

Fig. 4. Mean N1 magnitudes (Panel A), inter-trial phase coherence (Panel B) and theta power (Panel C). Mean N1 magnitude was obtained from 9 channels centered on FCz, between 85 and 135 ms post-stimulus onset. Phase coherence was quantified between 4 and 7 Hz, averaged across all scalp channels, between 60 and 160 ms after stimulus onset. Theta power was taken over a broad window (0–250 ms post onset), averaged across all channels, after a narrow bandpass filter was applied (see text for details). Error bars indicated standard error of the mean.

patients relative to controls. The picture was quite different during the repetitive standard stimuli, during which alpha band activity was the primary response across groups, and also differentiated between patients and controls. Controls showed greater alpha power and coherence than did patients, at approximately the same time as N1 amplitude peaks.

The primary distinction between a standard and deviant in the MMN protocol is the predictability of that stimulus. For the unpredictable stimulus, the spectral response was characterized by increases in theta power, but this response was reduced (or suppressed) for the repeated, or predictable, standard stimuli. This pattern is similar to (and consistent with) that seen in the current protocol using self-generated versus externally generated sounds: in the Listen condition, the externally generated sounds were less predictable for participants than were the self-generated sounds in the Talk condition, and elicited correspondingly greater theta power. On this account, one might therefore expect that further manipulations of predictability—by providing visual cues to enhance temporal predictability of externally generated sounds, or reducing predictability of content via pitch-shifting—would elicit further modulations of theta power and coherence. However, the effects of our temporal and content predictability manipulations were nuanced. No effect of temporal cueing (using an onscreen visual cue) or voice alteration was seen on N1 magnitude or theta power. However, both metrics impacted theta coherence. Visual cueing increased the predictability of the onset of sound stimuli, and led to an increase in theta coherence. Voice alteration reduced the predictability of the content of speech, and elicited reduced theta coherence. That is, theta coherence appeared to index the predictability of the sound stimuli in a manner that neither N1 magnitude nor theta power measures were able to. Note, however, that Cued and Listen conditions were not counterbalanced, so the cueing effect remains confounded with stimulus repetition (although, if anything, repetition of a standard stimulus tends to reduce event-related theta power, rather than increase it; Ko et al., 2012). Moreover, close inspection of responses to the standard stimuli in Lee et al.'s (2017) MMN task reveals a complementary effect: theta power was suppressed to the highly repeated (and thus predictable) standard stimuli, but theta coherence was increased for these stimuli (and more so in controls than in patients).

Superficially the effects of action and cueing on coherence in our data appear to contradict each other. Coherence was reduced when the resulting sound was predictable by virtue of the participant performing a sound-evoking action (speaking), whereas coherence was increased when the sound was predictable in its content (i.e. not modulated) or timing (by visual cueing). There is some precedent for this distinction between action and cueing when it comes to modulation of neural signatures of speech processing. Harrison et al. (2021) demonstrated separable effects of action- and cueing-based manipulations of predictability on N1 magnitude. Further, Jack et al. (2021) and Timm et al. (2014) found that the experience of performing a sound producing action (i.e. a sensory cue) was insufficient to elicit N1 suppression, and instead the participant must have intended to generate a sound in order for suppression to occur.

These distinct effects of action and cueing may be a consequence of the manipulations engaging different mechanisms. For example, the effect of action—here implemented as the difference between active talking and passive conditions—may reflect an influence of attention, as observed to occur across other active/passive EEG paradigms (Nobre, 2001; Hillyard, 1981). For example, actively listening to sounds (typically to detect a disruption in a sequence) results in different sensory potentials to passively listening to those same sounds (Horváth and Winkler, 2010). So the performance of an action (during the Talk condition) could draw attention away from listening to the sound, thereby influencing neural markers of stimulus processing (such as theta coherence or N1 amplitude) in a manner quite distinct from the manipulation of stimulus predictability via cueing. While there is little evidence to suggest that this kind of attentional explanation offers a full account of N1 suppression effects (see Baess et al., 2011), it may account for the differential impacts of action and cueing seen in the present coherence measure.

Notably, only theta coherence (specifically intertrial coherence) was affected by manipulations of stimulus predictability, in addition to the manipulation of self-generation. Although our tailored measure of theta power revealed significant sensory self-suppression (Talk vs Listen/Cued) of approximately the same effect size as coherence, power was insensitive to other measures of predictability. Cumulatively, this pattern lends further support to Roach et al.'s (2021) claim that theta coherence may be the superior metric of disruptions in sensory self-

suppression in schizophrenia, which are thought to be fundamentally elicited by the capacity to accurately predict (and then suppress) self-generated speech sounds. This conclusion accords with a broader literature that has observed theta coherence deficits in a broad range of tasks, including face emotion processing (Csukly et al., 2014) and in people's response to high-frequency, high contrast visual stimulus presentation (Hamilton et al., 2020). Perhaps more importantly, there is an increasing focus on the importance of theta coherence in the MMN literature, which itself has been proposed as a potential biomarker for schizophrenia (Javitt et al., 2000). Javitt et al. (2000) reviewed evidence of theta disruption in response to prediction error in humans and other animals, with particular reference to potential cortical disruptions that may explain the differential response in people with schizophrenia (or animal models of schizophrenia) relative to controls. They concluded that the impact of schizophrenia on theta frequency responses to prediction-error-evoking events (deviant stimuli in MMN, or potentially an externally generated or pitch-shifted sound in the present task) may be due to disruptions in the performance of SST interneurons that control superficial pyramidal neurons in the auditory cortex. These interneurons have been shown to have the functional characteristics needed to mediate an effect of stimulus predictability (Chen et al., 2015) and are also known to be disrupted in schizophrenia (Beneyto et al., 2012). It remains possible that this mechanism may extend to the effects of sensory self-suppression as well.

5. Conclusion

Sensory self-suppression was observed in healthy adults and people with schizophrenia in event-related potentials (N1 amplitude magnitude) and also in theta-band responses to speech stimuli (theta phase and coherence). Consistent with earlier observations, this suppression effect was generally reduced in people with schizophrenia, although these effects varied in the level of supporting evidence in the present data. Notably theta coherence was additionally affected by two manipulations of stimulus predictability (a visual cue indicating stimulus onset, and a voice alteration manipulation), such that theta coherence magnitude tended to mirror stimulus predictability. However, the direction of this effect appeared to reverse for increases in predictability brought about by direct control (i.e. performing speech) as distinct from predictions generated by the structure of the task itself (visual cueing or voice modulation). In sum, these findings provide further evidence that suppression (particularly theta coherence suppression) may have potential as a biomarker of schizophrenia.

CRediT authorship contribution statement

Oren Griffiths: Conceptualization, Formal analysis, Methodology, Software, Supervision, Project administration, Funding acquisition. **Bradley N. Jack:** Conceptualization, Methodology, Writing – review & editing. **Daniel Pearson:** Software, Methodology, Investigation. **Ruth Elijah:** Investigation. **Nathan Mifsud:** Investigation. **Nathan Han:** Investigation. **Sol Libesman:** Investigation. **Ana Rita Barreiros:** Data curation. **Luke Turnbull:** Data curation. **Ryan Balzan:** Writing – review & editing. **Mike Le Pelley:** Conceptualization, Methodology, Writing – review & editing, Funding acquisition. **Anthony Harris:** Supervision, Project administration, Funding acquisition. **Thomas J. Whitford:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Australian Research Council [DE150100667, DP140104394, DP200103288, DE220100739] and the National Health and Medical Research Council [APP1069487, APP2004067].

References

Andreou, C., Leicht, G., Nolte, G., Polomac, N., Moritz, S., Karow, A., Mulert, C., 2015. Resting-state theta-band connectivity and verbal memory in schizophrenia and in the high-risk state. Schizophr. Res. 161 (2–3), 299–307.
Baess, P., Horváth, J., Jacobsen, T., Schröger, E., 2011. Selective suppression of self-initiated sounds in an auditory stream: an ERP study. Psychophysiology 48 (9), 1276–1283.
Bansal, S., Ford, J.M., Spering, M., 2018. The function and failure of sensory predictions. Ann. N. Y. Acad. Sci. 1426 (1), 199–220. https://doi.org/10.1111/nyas.13686.
Basar-Eroglu, C., Schmiedt-Fehr, C., Marbach, S., Brand, A., Mathes, B., 2008. Altered oscillatory alpha and theta networks in schizophrenia. Brain Res. 1235, 143–152.
Behroozmand, R., Larson, C.R., 2011. Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. BMC Neurosci. 12 (1), 1–10.
Bendixen, A., SanMiguel, I., Schröger, E., 2012. Early electrophysiological indicators for predictive processing in audition: a review. Int. J. Psychophysiol. 83 (2), 120–131.
Beneyto, M., Morris, H.M., Rovensky, K.C., Lewis, D.A., 2012. Lamina-and cell-specific alterations in cortical somatostatin receptor 2 mRNA expression in schizophrenia. Neuropharmacology 62 (3), 1598–1605.
Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.M., Robbins, K.A., 2015. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. Front. Neuroinf. 9, 16.
Chen, I.W., Helmchen, F., Lütcke, H., 2015. Specific early and late oddball-evoked responses in excitatory and inhibitory neurons of mouse auditory cortex. J. Neurosci. 35 (36), 12560–12573.
Cohen, M.X., 2014. Analyzing Neural Time Series Data: Theory and Practice. MIT press.
Cohen, M.X., Gulbinaite, R., 2017. Rhythmic entrainment source separation: Optimizing analyses of neural responses to rhythmic sensory stimulation. Neuroimage 147, 43–56.
Csukly, G., Stefanics, G., Komlósi, S., Czigler, I., Czobor, P., 2014. Event-related theta synchronization predicts deficit in facial affect recognition in schizophrenia. J. Abnorm. Psychol. 123 (1), 178.
Delorme, A., Makeig, S., 2004. EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. Methods 134 (1), 9–21.
Duckworth, M., Allen, G., Ball, M., 1990. Extensions to the International Phonetic Alphabet for the transcription of atypical speech. Clin. Linguist. Phon. 4 (4), 273–280.
Eliades, S.J., Wang, X., 2003. Sensory-MOTOR INTERACTION IN THE PRIMATE AUDITORY CORTEX DURING SELF-INITIATED VOCALIZations. J. Neurophysiol. 89 (4), 2194–2207.
First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1997. Structured Clinical Interview for DSM-IV Axis I disorders (SCID I). Biometric Research Department, New York.
Fletcher, P.C., Frith, C.D., 2009. Perceiving is believing: a bayesian approach to explaining the positive symptoms of schizophrenia. Nat. Rev. Neurosci. 10 (1), 48–58.
Ford, J.M., Mathalon, D.H., Heinks, T., Kalba, S., Faustman, W.O., Roth, W.T., 2001. Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. Am. J. Psychiatry 158 (12), 2069–2071.
Ford, J.M., Mathalon, D.H., Kalba, S., Whitfield, S., Faustman, W.O., Roth, W.T., 2001. Cortical responsiveness during talking and listening in schizophrenia: an event-related brain potential study. Biol. Psychiatry 50 (7), 540–549.
Ford, J.M., Mathalon, D.H., Whitfield, S., Faustman, W.O., Roth, W.T., 2002. Reduced communication between frontal and temporal lobes during talking in schizophrenia. Biol. Psychiatry 51 (6), 485–492.
Ford, J.M., Mathalon, D.H., 2012. Anticipating the future: automatic prediction failures in schizophrenia. Int. J. Psychophysiol. 83 (2), 232–239.
Ford, J.M., Roach, B.J., Mathalon, D.H., 2010. Assessing corollary discharge in humans using noninvasive neurophysiological methods. Nat. Protoc. 5 (6), 1160–1168.
Ford, J.M., Mathalon, D.H., Roach, B.J., Keedy, S.K., Reilly, J.L., Gershon, E.S., Sweeney, J.A., 2013. Neurophysiological evidence of corollary discharge function during vocalization in psychotic patients and their nonpsychotic first-degree relatives. Schizophr. Bull. 39 (6), 1272–1280.
Fuentemilla, L., Marco-Pallarés, J., Grau, C., 2006. Modulation of spectral power and of phase resetting of EEG contributes differentially to the generation of auditory event-related potentials. Neuroimage 30 (3), 909–916.
Grau, C., Fuentemilla, L., Marco-Pallarés, J., 2007. Functional neural dynamics underlying auditory event-related N1 and N1 suppression response. Neuroimage 36 (3), 522–531.

Haddock, G., McCarron, J., Tarrier, N., Faragher, E.B., 1999. Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). Psychol. Med. 29 (4), 879–889.
Hamilton, H.K., Roach, B.J., Cavus, I., Teyler, T.J., Clapp, W.C., Ford, J.M., Mathalon, D. H., 2020. Impaired potentiation of theta oscillations during a visual cortical plasticity paradigm in individuals with schizophrenia. Front. Psychol. 11, 590567.
Harrison, A.W., Mannion, D.J., Jack, B.N., Griffiths, O., Hughes, G., Whitford, T.J., 2021. Sensory attenuation is modulated by the contrasting effects of predictability and control. Neuroimage 237, 118103.
Heinks-Maldonado, T.H., Mathalon, D.H., Houde, J.F., Gray, M., Faustman, W.O., Ford, J.M., 2007. Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. Arch. Gen. Psychiatry 64 (3), 286–296.
Hillyard, S.A., 1981. Selective auditory attention and early event-related potentials: a rejoinder. Can. J. Psychol. 35 (2), 159–174.
Horváth, J., Winkler, I., 2010. Distraction in a continuous-stimulation detection task. Biol. Psychol. 83 (3), 229–238.
Hughes, G., Desantis, A., Waszak, F., 2013. Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. Psychol. Bull. 139 (1), 133.
Jack, B.N., Chilver, M.R., Vickery, R.M., Birznieks, I., Krstanoska-Blazeska, K., Whitford, T.J., Griffiths, O., 2021. Movement planning determines sensory suppression: an event-related potential study. J. Cogn. Neurosci. 33 (12), 2427–2439.
JASP Team (2021). *JASP* (Version 0.16)[Computer software].
Javitt, D.C., Shelley, A.M., Ritter, W., 2000. Associated deficits in mismatch negativity generation and tone matching in schizophrenia. Clin. Neurophysiol. 111 (10), 1733–1737.
Kirkpatrick, B., Strauss, G.P., Nguyen, L., Fischer, B.A., Daniel, D.G., Cienfuegos, A., Marder, S.R., 2011. The brief negative symptom scale: psychometric properties. Schizophr. Bull. 37 (2), 300–305.
Kleiner, M., Brainard, D., Pelli, D., 2007. What's new in Psychtoolbox-3? Perception 36 (14), 1–16.
Ko, D., Kwon, S., Lee, G.T., Im, C.H., Kim, K.H., Jung, K.Y., 2012. Theta oscillation related to the auditory discrimination process in mismatch negativity: oddball versus control paradigm. J. Clin. Neurol. 8 (1), 35–42.
Kolev, V., Rosso, O.A., Yordanova, J., 2001. A transient dominance of theta ERP component characterizes passive auditory processing: evidence from a developmental study. Neuroreport 12 (13), 2791–2796.
König, T., Lehmann, D., Saito, N., Kuginuki, T., Kinoshita, T., Koukkou, M., 2001. Decreased functional connectivity of EEG theta-frequency activity in first-episode, neuroleptic-naıve patients with schizophrenia: preliminary results. Schizophr. Res. 50 (1–2), 55–60.
Lee, M., Sehatpour, P., Hoptman, M.J., Lakatos, P., Dias, E.C., Kantrowitz, J.T., Javitt, D. C., 2017. Neural mechanisms of mismatch negativity dysfunction in schizophrenia. Mol. Psychiatry 22 (11), 1585–1593.
Lovibond, P.F., Lovibond, S.H., 1995. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav. Res. Ther. 33 (3), 335–343.
Luck, S.J., Gaspelin, N., 2017. How to get statistically significant effects in any ERP experiment (and why you shouldn't). Psychophysiology 54 (1), 146–157.
Mathalon, D.H., Roach, B.J., Ferri, J.M., Loewy, R.L., Stuart, B.K., Perez, V.B., Ford, J.M., 2019. Deficient auditory predictive coding during vocalization in the psychosis risk syndrome and in early illness schizophrenia: the final expanded sample. Psychol. Med. 49 (11), 1897–1904.
MATLAB, V. (2010). 7. 10. 0 (R2010a). *The MathWorks Inc., Natick, Massachusetts*.
McGurn, B., Starr, J.M., Topfer, J.A., Pattie, A., Whiteman, M.C., Lemmon, H.A., Whalley, L.J., Deary, I.J., 2004. Pronunciation of irregular words is preserved in dementia, validating premorbid IQ estimation. Neurology 62 (7), 1184–1186.
Millard, S.J., Bearden, C.E., Karlsgodt, K.H., Sharpe, M.J., 2022. The prediction-error hypothesis of schizophrenia: New data point to circuit-specific changes in dopamine activity. Neuropsychopharmacology 47 (3), 628–640.
Mullen, T. (2012). NITRC: *CleanLine: Tool/Resource Info*. Available online at: http://www.nitrc.org/projects/cleanline (Accessed March Sept, 2021).
Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24 (4), 375–425.
Nasrallah, H., Morosini, P., Gagnon, D.D., 2008. Reliability, validity and ability to detect change of the Personal and Social Performance scale in patients with stable schizophrenia. Psychiatry Res. 161 (2), 213–224.
Nobre, A.C., 2001. Orienting attention to instants in ime. Neuropsychologia 39 (12), 1317–1328.
Oestreich, L.K., Mifsud, N.G., Ford, J.M., Roach, B.J., Mathalon, D.H., Whitford, T.J., 2015. Subnormal sensory attenuation to self-generated speech in schizotypy: electrophysiological evidence for a 'continuum of psychosis'. Int. J. Psychophysiol. 97 (2), 131–138.
Roach, B.J., Ford, J.M., Loewy, R.L., Stuart, B.K., Mathalon, D.H., 2021. Theta phase synchrony is sensitive to corollary discharge abnormalities in early illness schizophrenia but not in the psychosis risk syndrome. Schizophr. Bull. 47 (2), 415–423.
Roberts, T.P., Ferrari, P., Stufflebeam, S.M., Poeppel, D., 2000. Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. J. Clin. Neurophysiol. 17 (2), 114–129.
Sauseng, P., Klimesch, W., Gruber, W.R., Hanslmayr, S., Freunberger, R., Doppelmayr, M., 2007. Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. Neuroscience 146 (4), 1435–1444.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59 (20), 22–33.
Timm, J., SanMiguel, I., Saupe, K., Schröger, E., 2013. The N1-suppression effect for self-initiated sounds is independent of attention. BMC Neurosci. 14 (1), 1–11.
Timm, J., SanMiguel, I., Keil, J., Schröger, E., Schönwiesner, M., 2014. Motor intention determines sensory attenuation of brain responses to self-initiated sounds. J. Cogn. Neurosci. 26 (7), 1481–1489.
Uhlhaas, P.J., Singer, W., 2010. Abnormal neural oscillations and synchrony in schizophrenia. Nat. Rev. Neurosci. 11 (2), 100–113.
Van Wassenhove, V., Grant, K.W., Poeppel, D., 2005. Visual speech speeds up the neural processing of auditory speech. Proc. Natl. Acad. Sci. 102 (4), 1181–1186.
Wang, J., Mathalon, D.H., Roach, B.J., Reilly, J., Keedy, S.K., Sweeney, J.A., Ford, J.M., 2014. Action planning and predictive coding when speaking. Neuroimage 91, 91–98.
Whitford, T.J., 2019. Speaking-induced suppression of the auditory cortex in humans and its relevance to schizophrenia. Biol. Psychiatry Cognit. Neurosci. Neuroimag. 4 (9), 791–804.
Woods, D.L., 1995. The component structure of the N 1 wave of the human auditory evoked potential. Electroencephalogr. Clin. Neurophysiol. Suppl. Only 44, 102–109.
Yordanova, J., Kolev, V., 1998. Single-sweep analysis of the theta frequency band during an auditory oddball task. Psychophysiology 35 (1), 116–126.

Further reading

Chen, C.M.A., Mathalon, D.H., Roach, B.J., Cavus, I., Spencer, D.D., Ford, J.M., 2011. The corollary discharge in humans is related to synchronous neural oscillations. J. Cogn. Neurosci. 23 (10), 2892–2904.